

WHAT IS CLAIMED:

1. A composition comprising a purified proteolytic polypeptide derived from Hepatitis C virus.

2. The composition of claim 1, wherein said polypeptide has a partial internal sequence substantially as follows:

...Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr...

3. The composition of claim 1, wherein said polypeptide has a partial internal sequence substantially as follows:

...Leu Lys Gly Ser Ser Gly Gly Pro Leu...

4. The composition of claim 2, wherein said polypeptide has substantially the partial internal sequence:

Arg Arg Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Gly Met Val Ser Lys Gly
Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu
Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly Glu
Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys Ile Asn Gly
Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr¹ Ile Ala Ser Pro Lys
Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly Trp Pro
Ala Pro Gln Gly Ser Arg Ser Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu
Tyr Leu Val Thr Arg His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser
Arg Gly Ser Leu Leu Ser Pro Arg Pro Ile Ser Tyr² Leu Lys Gly Ser Ser Gly
Gly Pro Leu³ Leu Cys Pro Ala Gly His Ala Val Gly Ile Phe Arg Ala Ala Val
Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn Leu Glu
Thr Thr Met Arg.

5. The composition of claim 1, wherein said polypeptide has substantially the amino acid sequence shown in Figure 1.

- 41 -

6. A fusion protein, comprising:
a suitable fusion partner, fused to
a proteolytic polypeptide derived from Hepatitis C virus.

5 7. The fusion protein of claim 6, wherein said fusion partner comprises
human superoxide dismutase.

8. The fusion protein of claim 6, wherein said proteolytic polypeptide
has a partial internal sequence substantially as follows:
10 •••Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr•••.

9. The fusion protein of claim 6, wherein said proteolytic polypeptide
has a partial internal sequence substantially as follows:
•••Leu Lys Gly Ser Ser Gly Gly Pro Leu•••.

15 10. The fusion protein of claim 6, wherein said proteolytic polypeptide
has as a partial internal sequence:

20 Arg Arg Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Gly Met Val Ser
Lys Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr
Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys
Asn Gln Val Glu Gly Glu Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe
Leu Ala Thr Cys Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala
Gly Thr Arg Thr Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr
25 Asn Val Asp Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ser Arg
Ser Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr
Arg His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly
Ser Leu Leu Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly
Pro Leu Leu Cys Pro Ala Gly His Ala Val Gly Ile Phe Arg Ala Ala Val
30 Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn
Leu Glu Thr Thr Met Arg.

11. The fusion protein of claim 6, wherein said fusion partner is ubiquitin.

12. A composition comprising a polynucleotide which encodes only the HCV protease or an active HCV protease analog.

13. The composition of claim 12, wherein said polynucleotide encodes
5 the HCV protease of Figure 1.

14. A composition comprising a polynucleotide which encodes a fusion protein comprising:

10 HCV protease or HCV protease analog; and
a fusion partner.

15. The composition of claim 14, wherein said fusion partner is selected from the group consisting of hSOD, yeast α -factor, IL-2S, ubiquitin, β -galactosidase, β -lactamase, horseradish peroxidase, glucose oxidase, and urease.

16. The composition of claim 14, wherein said HCV protease or HCV protease analog comprises a polypeptide having substantially the following sequence:

20 Arg Arg Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Gly Met Val Ser
Lys Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr
Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys
Asn Gln Val Glu Gly Glu Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe
Leu Ala Thr Cys Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala
Gly Thr Arg Thr Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr
Asn Val Asp Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ser Arg
25 Ser Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr
Arg His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly
Ser Leu Leu Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly
Pro Leu Leu Cys Pro Ala Gly His Ala Val Gly Ile Phe Arg Ala Ala Val
Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn
30 Leu Glu Thr Thr Met Arg.

- 43 -

17. The composition of claim 14, wherein said HCV protease or analog comprises a polypeptide having substantially the sequence:

5 Gly Thr Tyr Val Tyr Asn His Leu Thr Pro Leu Arg Asp Trp Ala His
Asn Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe Ser
Gln Met Glu Thr Lys Leu Ile Thr Trp Gly Ala Asp Thr Ala Ala Cys
Gly Asp Ile Ile Asn Gly Leu Pro Val Ser Ala Arg Arg Gly Arg Glu Ile
Leu Leu Gly Pro Ala Asp Gly Met Val Ser Lys Gly Trp Arg Leu Leu
Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile
10 Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly Glu Val
Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys Ile Ile Asn
Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr Ile Ala Ser
Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp Gln Asp Leu
Val Gly Trp Pro Ala Ser Gln Gly Thr Arg Ser Leu Thr Pro Cys Thr
Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile
15 Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro Arg
Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu Leu Cys Pro Ala
Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys Thr Arg Gly Val
Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn Leu Glu Thr Thr Met
Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser
20 Phe Gln Val Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys
Val Pro Ala Ala.

18. The composition of claim 14, wherein said polypeptide has substantially the sequence:

25 Gly Thr Tyr Val Tyr Asn His Leu Thr Pro Leu Arg Asp Trp Ala His
Asn Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe Ser
Gln Met Glu Thr Lys Leu Ile Thr Trp Gly Ala Asp Thr Ala Ala Cys
Gly Asp Ile Ile Asn Gly Leu Pro Val Ser Ala Arg Arg Gly Arg Glu Ile
Leu Leu Gly Pro Ala Asp Gly Met Val Ser Lys Gly Trp Arg Leu Leu
30 Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile
Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly Glu Val
Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys Ile Ile Asn
Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr Ile Ala Ser
Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp Gln Asp Leu
35 Val Gly Trp Pro Ala Ser Gln Gly Thr Arg Ser Leu Thr Pro Cys Thr
Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile
Pro Val Arg.

3019201

- 44 -

19. A method for assaying compounds for activity against hepatitis C virus, comprising:

providing an active hepatitis C virus protease;

contacting said protease with a compound capable of inhibiting serine protease activity; and

measuring inhibition of the proteolytic activity of said hepatitis C virus protease.

20. An expression vector for producing HCV protease or HCV protease analogs in a host cell, which vector comprises:

a polynucleotide encoding HCV protease or an HCV analog;

transcriptional and translational regulatory sequences functional in said host cell operably linked to said HCV protease-encoding polynucleotide; and
a selectable marker.

21. The vector of claim 20, which further comprises a sequence encoding a fusion partner, linked to said HCV protease-encoding polynucleotide to form a fusion protein upon expression.

22. The vector of claim 21, wherein said fusion partner is selected from the group consisting of hSOD, yeast α -factor, IL-2S, ubiquitin, β -galactosidase, β -lactamase, horseradish peroxidase, glucose oxidase, and urease.

23. The vector of claim 22, wherein said fusion partner is selected from the group consisting of ubiquitin, hSOD, and yeast α -factor.

24. The vector of claim 20, wherein said HCV protease-encoding polynucleotide encodes a polypeptide having the substantially the following sequence:

3019202

- 45 -

Arg Arg Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Gly Met Val Ser
Lys Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr
Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys
Asn Gln Val Glu Gly Glu Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe
5 Leu Ala Thr Cys Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala
Gly Thr Arg Thr Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr
Asn Val Asp Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ser Arg
Ser Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr
Arg His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly
10 Ser Leu Leu Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly
Pro Leu Leu Cys Pro Ala Gly His Ala Val Gly Ile Phe Arg Ala Ala Val
Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn
Leu Glu Thr Thr Met Arg.

15 25. The vector of claim 20, wherein said HCV protease-encoding
polynucleotide encodes a polypeptide having the substantially the following sequence:

Gly Thr Tyr Val Tyr Asn His Leu Thr Pro Leu Arg Asp Trp Ala His
Asn Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe Ser
Gln Met Glu Thr Lys Leu Ile Thr Trp Gly Ala Asp Thr Ala Ala Cys
20 Gly Asp Ile Ile Asn Gly Leu Pro Val Ser Ala Arg Arg Gly Arg Glu Ile
Leu Leu Gly Pro Ala Asp Gly Met Val Ser Lys Gly Trp Arg Leu Leu
Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile
Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly Glu Val
Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys Ile Ile Asn
25 Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr Ile Ala Ser
Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp Gln Asp Leu
Val Gly Trp Pro Ala Ser Gln Gly Thr Arg Ser Leu Thr Pro Cys Thr
Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile
Pro Val Arg.

30

26. The vector of claim 20, wherein said HCV protease-encoding
polynucleotide encodes a polypeptide having the substantially the following sequence:

Gly Thr Tyr Val Tyr Asn His Leu Thr Pro Leu Arg Asp Trp Ala His
35 Asn Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe Ser
Gln Met Glu Thr Lys Leu Ile Thr Trp Gly Ala Asp Thr Ala Ala Cys
Gly Asp Ile Ile Asn Gly Leu Pro Val Ser Ala Arg Arg Gly Arg Glu Ile
Leu Leu Gly Pro Ala Asp Gly Met Val Ser Lys Gly Trp Arg Leu Leu

3019203

- 46 -

Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile
Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly Glu Val
Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys Ile Ile Asn
Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr Ile Ala Ser
5 Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp Gln Asp Leu
Val Gly Trp Pro Ala Ser Gln Gly Thr Arg Ser Leu Thr Pro Cys Thr
Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile
Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro Arg
Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu Leu Cys Pro Ala
10 Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys Thr Arg Gly Val
Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn Leu Glu Thr Thr Met
Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser
Phe Gln Val Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys
Val Pro Ala Ala.

3019204